Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biotechnology 119 (2005) 15–19
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

High through-put characterization of insect morphocryptic entities by a non-invasive method using direct-PCR of fecal DNA

B. Fumanal [a], J.-F. Martin [b], M.C. Bon [a,*]

[a] *European Biological Control Laboratory, USDA-ARS, 34980 Montferrier sur Lez, France*
[b] *Centre de Biologie et de Gestion des Populations, Ecole Nationale Supérieure Agronomique de Montpellier, Campus International de Baillarguet, 34980 Montferrier sur Lez, France*

Received 24 September 2004; received in revised form 30 March 2005; accepted 6 April 2005

Abstract

The development of non-invasive molecular techniques is currently increasing, particularly in the fields of behavioural ecology and conservation genetics of mammals. Surprisingly, genetic studies of Arthropods and particularly the insects have not benefited yet from the contributions that non-invasive methods have made. Here, we outline a strategy for identifying phytophagous insect genetic entities based on direct-PCR of fecal DNA combined with double strand conformation polymorphism (DSCP) typing. This allows the differentiation of morphocryptic entities within the species *Ceutorhynchus assimilis* (Coleoptera: Curculionidae), a candidate biocontrol agent of a noxious weed. The results obtained clearly demonstrate the potential for this method to provide a valuable means for genetic and ecological studies of Arthropods.
Published by Elsevier B.V.

Keywords: DSCP; Insect fecal DNA; Non-invasive method

1. Introduction

During the last several decades, the recognition of groups of insects, whether they are populations, species or higher taxonomic groups and, ultimately the understanding of the relationships between them, has fully benefited from the application of molecular methods to systematics, largely achieved by DNA sequencing. For the Arthropods, where the absence of morphological distinction prevails, accurate taxonomic identification becomes an issue, particularly in the fields of conservation genetics and biological control programs (Unruh and Wooley, 1999). At the heart of the problem is the fact that accurate molecular characterization primarily relies upon destructive methods, based upon DNA extracted from parts of or the entire insect (Unruh and Wooley, 1999), and gut contents in the case of predation studies (Symondson, 2002). This prevents any further biological study of the material analyzed. However, much of the recent developments and application of non-invasive techniques, based upon DNA

* Corresponding author. Fax: +33 499 623 049.
E-mail address: mcbon@ars-ebcl.org (M.C. Bon).

0168-1656/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.jbiotec.2005.04.011

studies have taken place in the area of dietary analysis of vertebrates. Regurgitated pellets from predatory birds (Taberlet and Fumagali, 1996) or mammal feces (Taberlet et al., 1997) provided the source material for DNA analyses. Such non-invasive diagnostic methods are now being harnessed for identifying and typing both prey and/or predators (Taberlet et al., 1997; Palomares et al., 2002).

This paper describes for the first time a non-invasive molecular method developed for the unequivocal identification of DNA from feces of a phytophagous weevil, *Ceutorhynchus assimilis* (Paykull, 1792) [*C. pleurostigma* (Marsham, 1802) syn.n.] (Coleoptera: Curculionidae) (Colonnelli, 1993) coupled with a reliable genetic typing of this species. We demonstrate that mitochondrial DNA including the COI region (Simon et al., 1994) directly acquired from fecal secretions (without preliminary DNA extraction) is of amplifiable quality comparable to that obtained from destructive methods. We also show the potential of the non-invasive technique to investigate the screening of the variation of this mitochondrial region by double strand conformation polymorphism typing (DSCP). DSCP is of great value for population genetics (Atkinson and Adams, 1997) and is detected when mutations alter the curvature of the helical axis of double stranded DNA molecules, resulting in changes in their electrophoretic mobility (Saad et al., 1994).

C. assimilis is a potential biocontrol agent of a noxious weed, *Lepidium draba* spp. *draba* (Brassicaceae), invading the USA. Fumanal et al. (2004a) showed that this recognized species is in fact a complex of generalist entities with one entity restricted to the host *L. draba* over its Eurasian distribution range, all forms being morphologically indistinguishable. Outcrossing testing was selected at first in order to delineate the taxonomical range of these forms according to the prevalent biological species concept.

2. Materials and methods

Four parental populations of *C. assimilis* were chosen based on multiple criteria inferred from both host-specificity testing (Fumanal et al., 2004a) and DNA analysis results (Fumanal et al., 2004b). In this last study, the mitochondrial haplotypic diversity of *C. assimilis* was primarily assessed based on sequencing data of a 560 bp region in the cytochrome oxydase I gene (COI) using the primer pair C1-J-2183 and TL2-N-3014 (Simon et al., 1994). The results showed that these four populations clustered in three mtDNA haplotypic groups (clades). Individuals of the population A, which originated from the target weed, *L. draba* spp. *draba*, in France, were shared between the clade II (host-specific entities) and clade IV (generalist entities). Both populations B and C, respectively, collected on the target weed in Italy and on *Brassica napus* in France, were belonging to the clade IV. The clade III was gathering generalist entities of the population D collected from another host-plant *Sinapis arvensis*, in France.

Mature larvae of *C. assimilis* developing in collar galls in the host-plant were collected in the field, and extracted in the laboratory or in quarantine facilities for foreign biological material. One month after pupation, the emerging immature adults were sex-determined, kept isolated and fed with fresh leaf of their natural host-plant. After this three-day diet, feces were collected. In the absence of feces, adults were fed with fresh leaves of *L. draba* exclusively for one to three more days. Then, two dried feces of each individual were collected under a binocular microscope using pins and under sterile conditions to avoid contamination from other insects, and were suspended in 5 μL of pure water. To detect whether contamination of samples of exogenous DNA had occurred during the feces extraction procedure, a tube without feces (negative control) was treated identically through the amplification procedure.

An aliquot of 2 μL of this suspension was used as template for direct-PCR using primer pair C1-J-2183 (Simon et al., 1994) and C1-N-2471 (after nomenclature in Simon et al., 1994) [5′-GAAAAGTTG-CTAATCATCTG-3′]. This primer was designed for this study to amplify a 288 bp fragment of the COI gene using the Oligo 4.0 Software (National Biosciences Inc.) (Fig. 1). The selection of this region was a two-step procedure. First, the three common haplotypes (560 bp region in COI gene), found in the parental populations were aligned using Clustal X program version 1.81 (Thompson et al., 1997) to locate sites that varied among the three clades. Second, predicted curvature was plotted along the sequences using the Bend-it software as developed by Munteanu et al. (1998). This resulted in the identification of the region that displayed

```
                C1-J-2183
Clade II    CAACATTTATTTTGATTTTTTGGACATCCAGAAGTATATATTTTAATTTT   50
Clade III   ··················································
Clade IV    ··················································

Clade II    ACCTGGATTTGGGATAATTTCTCATATTATTGTTCAAGAAAGAGGAAAAA   100
Clade III   ··················································
Clade IV    ··················································

Clade II    AAGAAACCTTTGGTGTTTTAGGAATAATTTATGCTATAATAGCTATTGGA   150
Clade III   ······················T···························
Clade IV    ······················T···························

Clade II    TTACTAGGATTTGTAGTATGAGCTCATCATATATTTACAGTTGGAATAGA   200
Clade III   ············································A·····
Clade IV    ············································G·····

Clade II    CGTAGACACTCGTGCTTATTTCACATCAGCAACTATAATTATTGCAGTAC   250
Clade III   T·······················T·························
Clade IV    T·······················C·························
                                   C1-N-2471
Clade II    CAACAGGAATTAAAATTTTTAGATGATTAGCAACTTTTCATGGGGCCCAA   300
Clade III   ····C··············C······C·······················
Clade IV    ···················C······························
```

Fig. 1. Alignment of 5′–3′ sequences (300 bp region in the COI gene) of the three clades. Dots indicate identity. PCR primers (C1-J-2183 and C1-N-2471) for clade diagnostic are indicated in bold.

sufficient DSCP between the different entities to be diagnostic. Double-stranded DNA amplifications were performed in a 25 μL volume containing 10 μM of each primer, 2.5 mM dNTP, 1× Qiagen Buffer, 1 unit of Qiagen Taq Polymerase (5 U/μL) and 2 μL of template. Feces in solution in 2 μL of water were added to this mix in a 0.2 mL PCR tube. Thermal cycle parameters were as follows: 92 °C for 1 min followed by 35 cycles of 92 °C for 30 s, 48 °C for 1 min and 62 °C for 30 s. The last elongation step was extended to 7 min. PCR products were checked on 1% agarose gel electrophoresis (Fig. 2a). After recalibration to ensure similar quantities of DNA in all wells, PCR products were separated by electrophoresis on 10% polyacrylamide 49:1 gel (30% acrylamide, 5% glycerol) in TBE 1× buffer, for 15 h at 100 V, 30 mA in 20 cm × 20 cm plates. Gels

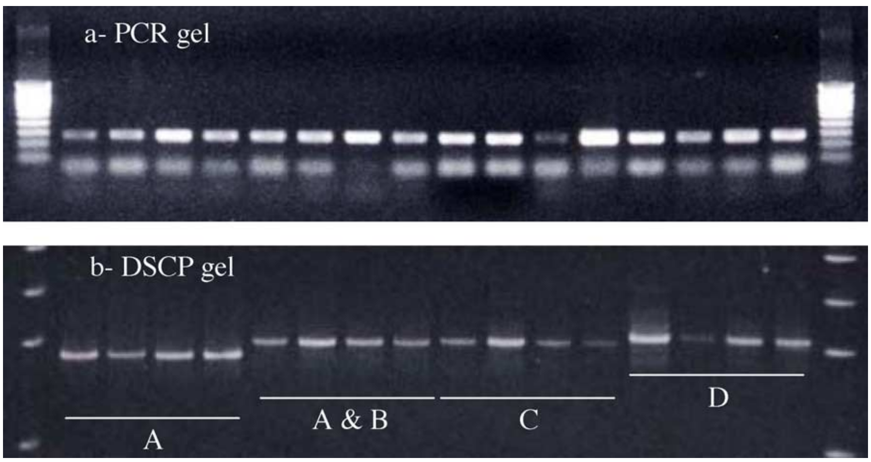

Fig. 2. (a) Direct-PCR products (PCR mix plus feces) of 317 bp checked on 1% agarose gel electrophoresis and (b) DSCP polyacrylamide gel (10%, 49:1) with direct-PCR products displaying band polymorphism corresponding to the different weevil populations (A–D).

were visualized using Ethidium bromide coloration (302 nm UV light) (Fig. 2b). Three distinct migration profiles were scored and have been ascribed to standards for known DNA sequences (Fig. 2b).

3. Results and discussion

PCR products of 288 bp were successfully amplified from feces of the weevil. The sequence produced from these PCR products corresponded to the expected ones. Details of the success rates for both molecular steps of the method (direct-PCR and DSCP) are presented in Table 1. The overall mean success obtained (all steps included) is 71.5% (448 positive results/626 attempts). Amplification of the 288-bp fragment from the fecal samples occurred successfully in population A–C. In the case of population D, PCR amplifiability averaged 57.3% only. One possible explanation for this lower score is that population D, which emerged first, was probably hampered by the fact that the protocol was not yet optimized. Amplification of a larger mtDNA fragment of 800 bp did not perform well unlike the amplification of the 288 bp fragment (data not shown). The authors' caution that PCR amplifiability of feces stored at 4 °C or room temperature were best preserved compared to those stored at −20 °C.

The failure to get DSCP identification after a successful PCR amplification was rather low, ranging from 2.6% (population A) to 16.5% (population B) (Table 1). This failure is not critical given that it is very easy to re-amplify the same individual using other secretions.

In the context of outcrossing *C. assimilis* populations, direct-PCR followed by the DSCP technique allowed successful discrimination between the three genetic entities (Fig. 2b). Populations B–D were monotypic with regard to the DSCP pattern, diagnostic of the generalist entities (Fig. 2b) whereas two different patterns were observed for population A, indicating a mixture of different genetic entities, including the host-specific one. Indeed, 90% of fecal samples belonged to the host restricted form, as expected and further selected for interbreeding tests, whereas 9% were displaying the B type DSCP pattern. The remaining 1% was undetermined, because of dubious multiple bands. The technique developed here proved to be very useful for untangling the mixture of different genetic clades found in the field and ensured reliable identification for subsequent interbreeding tests.

The procedure described above fills an empty niche in the field of non-invasive identification methods for which an array of tools has become increasingly available including recently developed molecular genetic techniques for the analysis of feces (Taberlet and Fumagali, 1996; Symondson, 2002). The *Ceutorhynchus* mtDNA region is not altered by digestive damage to diagnostic features. The failure to repeatedly amplify a 800 bp fragment when it is possible to amplify shorter fragments might reflect the fact that the DNA embedded in the epithelial cells could sometimes be partially degraded. This result strongly supports the assumption that DNA would not survive the digestive process for long, explaining why the introduction of this approach has been delayed. Several specialists of the non-invasive approaches suggest that larger sequences simply break down more rapidly during digestion than shorter sequences and that single-copy DNA would be less easily detected than multiple-copy DNA (Taberlet and Fumagali, 1996; Symondson, 2002). We chose to target mtDNA with 100–1000 copies per cell (Simon et al., 1994) this having been effectively used in other studies to look at fine-scale differences between closely related species or subspecies The amount of DNA available for genetic typing using feces can be very low and is often in the picogram range (Taberlet et al., 1996). In the case of *C. assimilis*, we used DSCP, which did not require extensive optimization and has proven to

Table 1
Success rates of the four population samples studied, for both direct-PCR and DSCP steps are presented

Population	No. of individuals	PCR success (%)	DSCP success (%)	Total PCR + DSCP (%)
A	211	91.5	97.4	89.1
B	131	83.2	83.5	69.5
C	120	78.3	94.7	74.2
D	164	57.3	92.6	53

Number of individual samples analyzed were added together with success rates in % for PCR, DSCP and PCR plus DSCP.

be of value for detecting sequence variation between morphocryptic entities but other post-PCR techniques could be used (SSCP or direct sequencing for instance).

In mammals, the intestinal lining cells shed in the feces have been shown to be an effective means of acquiring DNA as well as the digestibility of the diet to influence the digestion time, and thus to decrease the abrasion of the intestinal membranes (Taberlet et al., 1996; Maudet et al., 2004). However, the primary mechanism underlying the release of DNA in the feces still remains speculative for the insects. The authors' caution, that the *L. draba* like other Brassicaceae plants which is used as diet contains plant defense chemicals including glucosinolates (Fumanal et al., 2004a,b). Most known mechanisms used by insects to overcome these compounds involve detoxification and excretion (Wittstock et al., 2004). However, their role on the release of cells in the feces have yet to be determined.

In conclusion, direct-PCR combined with typing methods of fecal DNA fragments opens new perspectives for insect identification both in applied and basic science. Typing data using non-invasive methods, which would be generated from voucher specimens could easily be transferred to the growing databases that are being built up at national museums or other collections. Quarantine interception facilities and biological control agencies would also greatly benefit from this rapid identification method that reduces time consuming breeding procedures.

Acknowledgements

This work was supported in part by H. McNeel, Bureau of Land Management (BLM). We thank Réjane Streiff (CBGP, France), Rouhollah Sobhian, Corinne Hurard, Alan Kirk and William Meikle (EBCL) for helpful comments on the manuscript.

References

Atkinson, L., Adams, E.S., 1997. Double-strand conformation polymorphism (DSCP) analysis of the mitochondrial control generates highly variable markers for population studies in a social insect. Insect Mol. Biol. 6, 369–376.

Colonnelli, E., 1993. The Ceutorhynchinae types of I.C. Fabricius and G. von Paykull (Coleoptera: Curculionidae). Koleopterologishe Rundschau 63, 299–310.

Fumanal, B., Martin, J.F., Sobhian, R., Blanchet, A., Bon, M.C., 2004a. Host range of *Ceutorhynchus assimilis* (Coleoptera: Curculionidae) a candidate for biological control of *Lepidium draba* (Brassicaceae) in the USA. Biol. Control 30 (3), 598–607.

Fumanal, B., Martin, J.F., Sobhian, R., Gaskin, J., Bon, M.C., 2004b. Evolutionary biology as a tool towards a more customized biological control strategy of weeds: *Lepidium draba* as a case study. In: AFPP (Eds.), XIIth International Symposium on Weeds Biology, Dijon, France, pp. 421–426.

Maudet, C., Luikart, D., Dubray, D., Von Hardenberg, A., Taberlet, P., 2004. Low genotyping error rates in wild ungulate faeces sampled in winter. Mol. Ecol. Notes 4, 772–775.

Munteanu, M.G., Vlahovicek, K., Parthasaraty, S., Simon, I., Pongor, S., 1998. Rod models of DNA: sequence-dependent anisotropic elastic modelling of local bending phenomena. Trends Biochem. Sci. 23, 341–346.

Palomares, F., Godoy, J.A., Piriz, A., O'Brien, S.J., Johnson, W.E., 2002. Faecal genetic analysis to determine the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. Mol. Ecol. 11, 2171–2182.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25 (24), 4876–4882.

Saad, F.A., Halliger, B., Müller, C.R., Roberts, R.G., Danieli, G.A., 1994. Single base substitutions are detected by double strand conformation analysis. Nucleic Acids Res. 22 (20), 4352–4353.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Entomol. Soc. Am. 87, 651–701.

Symondson, W.O.C., 2002. Molecular identification of prey in predator diets. Mol. Ecol. 11, 627–641.

Taberlet, P., Fumagali, L., 1996. Owl pellets as source of DNA for genetic studies of small mammals. Mol. Ecol. 5, 301–305.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P., Bouvet, J., 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acid Res. 24 (16), 3189–3194.

Taberlet, P., Camarra, J.J., Griffin, S., Uhres, E., Hanotte, O., Waits, L.P., Dubois-Paganon, C., Burke, T., Bouvet, J., 1997. Noninvasive genetic tracking of the endangered Pyrenean brown bear population. Mol. Ecol. 6, 869–876.

Unruh, T., Wooley, J.B., 1999. Molecular methods in classical biological control. In: Belows, T.S., Fisher, T.W. (Eds.), Handbook of Biological Control—Principles and Applications of Biological Control. Academic Press, San Diego, pp. 57–85.

Wittstock, U., Agerbirk, N., Stauber, E., Olsen, C.E., Hippler, M., Mitchell-Olds, T., Gershenzon, J., Vogel, H., 2004. Successful herbivore attack due to metabolic diversion of a plant chemical defense. P.N.A.S. 101 (14), 4859–4864.